CAT + CBRUG R/RStudio workshop agenda

- **10am to 12pm**
- CAT to help/support learners with install and getting software open
- Intro and overview of R/RStudio
- Create a project and get data
  - Pull in Census2020 data from csv, excel, and from file pane
- Exploratory Data Analysis, commands to inspect dataframe
  - Str (structure)
  - Glimpse
  - Dim (dimensions)
  - Head/tail
  - Colnames/names
  - Min/max
  - Summary
  - View
  - Isolate a column with $
  - Indexing- df[row,col], run through a few different combinations
  - Assignment operator
- dplyr commands
  - Load packages
    - Read excel
    - Tidyverse
    - tidycensus
  - Get data: 2020 Census, 2019 ACS and 2019 Poverty
    - Excel, from drive
    - tidycensus
  - dplyr::select
    - keep variables
    - drop variables
  - dplyr::rename
    - rename columns
    - bonus1: chain select and rename
  - **BREAK**
  - dplyr::filter
    - filter by size
    - filter using and/or
  - dplyr::arrange
    - arrange ascending by size
    - arrange descending by size

- ▪ bonus2: chain filter and arrange
  - o dplyr::mutate
    - ▪ calculate and create a 2010 column
  - o dplyr::summarize
    - ▪ summarize the 2020 and 2010 population
    - ▪ bonus3: calculate mean pop in 2020 and 2010
    - ▪ bonus4: calculate the different of sum and mean between 2020 and 2010
    - ▪ bonus5: calculate the sum of large states
    - ▪ bonus6: calculate the sum of small states
  - o dplyr::group_by
    - ▪ create a change direction column, group_by the column and summarize the two groups
    - ▪ bonus7: chain mutate, group_by and summarize
- **BREAK for lunch**
- **1pm to 3pm**
  - o Combining data
    - ▪ join() two datasets- 2020 Census Pop and 2019 ACS Poverty
    - ▪ cbind()
    - ▪ rbind() example
  - o mutate() using new combined dataset to create a ranking variable and filter() the top 10
  - o Visualization using ggplot()
    - ▪ basic visualization from previous dataset, bar chart
    - ▪ histogram of 2020 state population
    - ▪ box plot of 2020 population by state size
      - ▪ use mutate to create movement direction, size bin and size rank
    - ▪ bar chart of 2020 population
    - ▪ scatterplot of 2020 population
    - ▪ scatterplot of 2019 population and poverty status with regression
    - ▪ bar chart faceted by size_bin
  - o **BREAK**
  - o Markdown demonstration
  - o Github demonstration
  - o Extra credit assignment
  - o Wrap up, QA